

Tracking Dynamics of Two-Dimensional Continuous Attractor Neural Networks

C. C. Alan Fung¹, K. Y. Michael Wong¹ and Si Wu²

¹Department of Physics, The Hong Kong University of Science and Technology, Hong Kong, China

²Lab of Neural Information Processing, Institute of Neuroscience, Shanghai, China

E-mail: alanfung@ust.hk, phkywong@ust.hk, siwu@ion.ac.cn

Abstract. We introduce an analytically solvable model of two-dimensional continuous attractor neural networks (CANNs). The synaptic input and the neuronal response form Gaussian bumps in the absence of external stimuli, and enable the network to track external stimuli by its translational displacement in the two-dimensional space. Basis functions of the two-dimensional quantum harmonic oscillator in polar coordinates are introduced to describe the distortion modes of the Gaussian bump. The perturbative method is applied to analyze its dynamics. Testing the method by considering the network behavior when the external stimulus abruptly changes its position, we obtain results of the reaction time and the amplitudes of various distortion modes, with excellent agreement with simulation results.

1. Introduction

Continuous attractor neural networks (CANNs) are very useful models for describing the encoding of continuous stimuli in neural systems [1–5]. The encoded stimuli can either be some simple features of objects, such as their orientations [6], moving directions [7] and spatial locations [8], or some complicated rules that underly the categorization of objects [9]. Compared with other attractor models, CANNs have the distinctive feature that they hold a family of stationary states which can be translated into each other without the need to overcome energy barriers. In the continuum limit, these stationary states form a continuous manifold in which the system is neutrally stable, and the network state can translate easily when the external stimulus changes continuously. Beyond pure memory retrieval, this large-scale structure of the state space endows the neural system with a tracking capacity.

To construct a model for CANN, the key is that the neuronal interactions should be properly balanced in excitation and inhibition (e.g., of the Mexican-hat shape) and be translationally invariant. The former enables the network to have a local persistent bump solution and the latter ensures that the network has a continuous family of such solutions. Although mathematically it is possible to construct a CANN of dimensionality larger than two, the research on CANNs has so far been mainly focused on one or two dimensional case. This is because in the cortex, neurons are essentially distributed in a two-dimensional sheet. To maintain a CANN of dimensionality larger than two, it is difficult to wire neurons in a two-dimensional space without their signals interfering with each other. To encode continuous features of high dimensionality, the brain may employ layers of neurons to combine low-dimensional CANNs hierarchically.

The tracking dynamics of a CANN has been theoretically investigated by several authors in the literature [5, 6, 10–12]. These studies have demonstrated that a CANN has the capacity of tracking a moving stimulus continuously and that this tracking behavior can describe many brain functions well. Despite these successes, however, detailed rigorous analysis of tracking behaviors of a CANN is still lacking. In a recent work [13, 14], the authors have developed a perturbative approach to elucidate the tracking performance of a one-dimensional CANN clearly. Because of the bump shape of the stationary states of the network, we used the wave functions of the quantum harmonic oscillator as the basis to decompose the network dynamics into different motion modes. These modes have clear physical meanings, corresponding to distortions in the amplitude, position, width or skewness of the network state. Due to the neutral stability of network states, the dynamics of a CANN is typically dominated by a few motion modes, with their contributions determined by the corresponding eigenvalues. We therefore can project the network dynamics on these dominating modes and simplify the network dynamics significantly. In this study, we extend the perturbative approach to a two-dimensional CANN. The two-dimensional CANN has much richer dynamics and distortion patterns than the one-dimensional case [15–17]. To elucidate the effect of different distortion patterns on the network dynamics clearly, we develop the perturbative approach in both rectangular and polar coordinates. To test our method, we study the tracking performance of the network when the external stimulus position experiences an abrupt change. Simulation results confirm that our method works very well.

2. The Model

We consider a two-dimensional neural network coding the stimulus $\mathbf{x} = (x_1, x_2)$, with N neurons distributed over this space. For simplicity, the neurons are assumed to be uniformly distributed in the space. Considering the common case that the range of possible values of the stimulus being much larger than the range of neuronal interactions, we can effectively take $-\infty < x_1, x_2 < \infty$. The dynamics of the synaptic input $U(\mathbf{x}, t)$ and neuronal response $r(\mathbf{x}, t)$ is given by

$$\tau \frac{\partial U(\mathbf{x}, t)}{\partial t} = I^{\text{ext}}(\mathbf{x}) + \rho \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{x}' J(\mathbf{x}, \mathbf{x}') r(\mathbf{x}') - U(\mathbf{x}, t); \quad (1)$$

$$r(\mathbf{x}, t) = \frac{U(\mathbf{x}, t)^2}{1 + k\rho \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{x}' U(\mathbf{x}', t)^2}, \quad (2)$$

where $J(\mathbf{x}, \mathbf{x}')$ is the translationally invariant coupling function defined by

$$J(\mathbf{x}, \mathbf{x}') = \frac{A}{2\pi a^2} \exp \left[-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2a^2} \right], \quad (3)$$

a is the tuning width of the neural network, k is the global inhibition, and ρ is the density of neurons over the space. When $I^{\text{ext}} = 0$ and $0 < k < k_c \equiv A^2\rho/(32\pi a^2)$, we have the steady solutions, or *stationary states*, given by (see figure 1)

$$\tilde{U}(\mathbf{x}|\mathbf{z}) = U_0 \exp \left[-\frac{|\mathbf{x} - \mathbf{z}|^2}{4a^2} \right], \quad (4)$$

$$\tilde{r}(\mathbf{x}|\mathbf{z}) = r_0 \exp \left[-\frac{|\mathbf{x} - \mathbf{z}|^2}{2a^2} \right], \quad (5)$$

where $U_0 = [1 + (1 - k/k_c)^{1/2}]A/(8\pi a^2 k)$ and $r_0 = [1 + (1 - k/k_c)^{1/2}]/(4\pi a^2 k\rho)$. It is notable that Eqs. (4) and (5) are valid for any \mathbf{z} . For simplicity, we consider $I^{\text{ext}} = \alpha U_0 \exp[-|\mathbf{x} - \mathbf{z}_0|^2/(4a^2)]$, where α is the strength of the stimulus. Thanks to the translational invariance of the coupling

function, the stationary state solution can be peaked at any point in the space. In this paper, we consider the network response to a stimulus abruptly changed from $(x_1, x_2) = (0, 0)$ to (z_{01}, z_{02}) at $t = 0$. As shown in the simulation result in figure 2, the synaptic input can track the change.

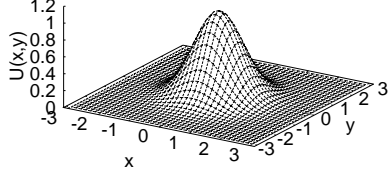


Figure 1. A stationary state $\tilde{U}(\mathbf{x}, 0)$. Parameters: $N = 40 \times 40$, $a = 0.5$, $k = 0.5$, $\tau = 1$ and $\rho = N/(2\pi)^2$.

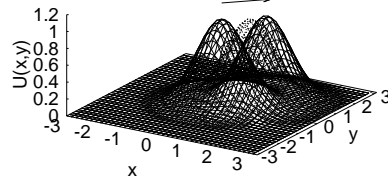


Figure 2. The synaptic input in the space due to the abruptly changed stimulus. Parameters: $\alpha = 0.05$, $\mathbf{z}_0(t < 0) = 0$, $\mathbf{z}_0(t \geq 0) = (1, 1)$ and the rest are the same as figure 1.

3. Solution to the Model

Under the driving of an external stimulus, the network state (i.e, the bump) moves from its initial position to the target one, with its shape distorted during the tracking process. Thus, to describe the tracking performance of a CANN, the key is to know the distortion patterns and their effects on the network dynamics. We denote the the state distortion to be $\delta U(\mathbf{x}, t) \equiv U(\mathbf{x}, t) - \tilde{U}(\mathbf{x}|\mathbf{z})$, whose dynamics is given by linearizing Eq.(1) at $\tilde{U}(\mathbf{x}|\mathbf{z})$ [13, 14],

$$\tau \frac{\partial}{\partial t} \delta U(\mathbf{x}, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{x}' F(\mathbf{x}, \mathbf{x}'|\mathbf{z}) \delta U(\mathbf{x}', t) - \delta U(\mathbf{x}, t), \quad (6)$$

where the interaction kernel $F(\mathbf{x}, \mathbf{x}'|\mathbf{z})$ is

$$F(\mathbf{x}, \mathbf{x}'|\mathbf{z}) = \frac{2\rho U(\mathbf{x}')}{B} \left[J(\mathbf{x}, \mathbf{x}') - k\rho \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{x}'' J(\mathbf{x}, \mathbf{x}'') r(\mathbf{x}'') \right], \quad (7)$$

where $B = 1 + k\rho \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{U}(\mathbf{x}'|\mathbf{z})^2 d\mathbf{x}'$ is a constant. The network dynamics is determined by the eigenfunctions and eigenvalues of the kernel $F(\mathbf{x}, \mathbf{x}'|\mathbf{z})$. To compute them, we choose the eigenfunctions of the quantum harmonic oscillator as the basis. While the results are more clearly presented using basis functions in polar coordinates, the analysis is more conveniently done in the rectangular coordinates. Hence we will describe the basis functions in both coordinates.

3.1. Using Basis Functions in Rectangular Coordinates

Under the rectangular coordinates, the basis functions are

$$u_{n,m}(\mathbf{x}|\mathbf{z}) = \sqrt{\frac{1}{2\pi a^2 m! n! 2^{n+m}}} H_n \left(\frac{x_1 - z_1}{\sqrt{2}a} \right) H_m \left(\frac{x_2 - z_2}{\sqrt{2}a} \right) \exp \left[-\frac{|\mathbf{x} - \mathbf{z}|^2}{4a^2} \right], \quad (8)$$

where $H_n(x)$ is the n^{th} order Hermite Polynomial [18]. By the completeness of the basis functions, we have

$$U(\mathbf{x}, t) = \tilde{U}(\mathbf{x}|\mathbf{z}(t)) + \sum_{n,m} a_{n,m}(t) u_{n,m}(\mathbf{x}|\mathbf{z}(t)). \quad (9)$$

Thanks to the orthonormality of the basis functions, we have derived [13, 14]

$$\begin{aligned} \tau \frac{da_{n,m}}{dt} &= I_{n,m} + \sum_{n',m'} \mathcal{F}_{n,m,n',m'} a_{n',m'}(t) - a_{n,m} \\ &\quad - \frac{\tau}{2a} \frac{dz_1}{dt} \left[\sqrt{2\pi} a U_0 \delta_{n1,m0} + \sqrt{n} a_{n-1,m}(t) - \sqrt{n+1} a_{n+1,m}(t) \right] \\ &\quad - \frac{\tau}{2a} \frac{dz_2}{dt} \left[\sqrt{2\pi} a U_0 \delta_{n0,m1} + \sqrt{m} a_{n,m-1}(t) - \sqrt{m+1} a_{n,m+1}(t) \right], \end{aligned} \quad (10)$$

where $I_{n,m}$ is the projection of I^{ext} onto $u_{n,m}$, given by

$$I_{n,m}(\mathbf{x}|\mathbf{z}_0) = \alpha U_0 \sqrt{\frac{2\pi a^2}{n!m!}} \left(\frac{z_{01} - z_1}{2a} \right)^n \left(\frac{z_{02} - z_2}{2a} \right)^m \exp \left[-\frac{|\mathbf{z} - \mathbf{z}_0|^2}{8a^2} \right], \quad (11)$$

and $\mathcal{F}_{n,m,n',m'}$, the interaction matrix, is given by

$$\mathcal{F}_{0,0,0,0} = 1 - \left(1 - \frac{k}{k_c} \right)^{\frac{1}{2}}, \quad (12)$$

$$\begin{aligned} \mathcal{F}_{n,m,n',m'} &= \sqrt{\frac{m'!n'!}{m!n!}} \frac{2}{2^{m'+n'}} \frac{(-)^{\frac{m'+n'-m-n}{2}}}{2^{\frac{m'+n'-m-n}{2}} (\frac{m'-m}{2})! (\frac{n'-n}{2})!}, \\ &\quad \text{if } \frac{n'-n}{2} \text{ and } \frac{m'-m}{2} \text{ are positive integers} \end{aligned} \quad (13)$$

$$\mathcal{F}_{n,m,n',m'} = 0, \quad \text{otherwise.} \quad (14)$$

The center of mass is given by the self-consistent condition $\mathbf{z}(t) = \iint d\mathbf{x} U(\mathbf{x}, t) \mathbf{x} / \iint d\mathbf{x} U(\mathbf{x}, t)$. If the external stimulus is symmetric with respect to the x -axis, then $dz_2/dt = 0$ and

$$\frac{dz_1}{dt} = \frac{2a}{\tau} \frac{\sum_{\text{odd } n, \text{ even } m} \sqrt{\frac{(m-1)!!}{m!!}} \frac{n!!}{(n-1)!!} \left(I_{n,m} + \sum_{n',m'} \mathcal{F}_{n,m,n',m'} a_{n',m'} \right)}{\sqrt{2\pi} a U_0 + \sum_{\text{even } n, m} \sqrt{\frac{(m-1)!!}{m!!}} \frac{(n-1)!!}{n!!} a_{n,m}}. \quad (15)$$

The eigenvalues of \mathcal{F} are $\lambda_{0,0} = 1 - (1 - k/k_c)^{1/2}$, $\lambda_{n,0} = \lambda_n$ for $n \neq 0$, $\lambda_{0,m} = \lambda_m$ for $m \neq 0$, $\lambda_{n,m} = \lambda_n \lambda_m$ for $n \neq 0$ and $m \neq 0$, where $\lambda_n = 2/2^n$. From this result, one can conclude that, if the stimulus is absent, all modes of distortion will decay exponentially in time, except for the eigenfunctions $u_{1,0}$ and $u_{0,1}$, whose eigenvalues are 1. To prove this, we define $v_{n,m}^R(\mathbf{x}|\mathbf{z})$ to be the right eigenfunctions of \mathcal{F} . Then we may express $\delta U(\mathbf{x}, t)$ in Eq. (6) as $\delta U(\mathbf{x}, t) = \sum_{n,m} \delta U_{n,m}(\mathbf{z}, t) v_{n,m}^R(\mathbf{x}|\mathbf{z})$. Using the orthonormality of the left and right eigenfunctions of $F(\mathbf{x}, \mathbf{x}'|\mathbf{z})$, we have

$$\delta U_{n,m}(z, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{x} \delta U(\mathbf{x}, t) v_{n,m}^L(\mathbf{x}|\mathbf{z}). \quad (16)$$

Assume that the motion of the bump is slow, so that $d\mathbf{z}/dt$ becomes negligible in Eq. (6); as we shall see, this assumption is valid as long as the external stimulus is sufficiently weak. Then, the projection of Eq. (6) on the eigenfunctions become

$$\tau \frac{d}{dt} \delta U_{n,m}(\mathbf{z}, t) = (\lambda_{n,m} - 1) \delta U_{n,m}(z, t). \quad (17)$$

Hence,

$$\delta U_{n,m}(\mathbf{z}, t) = \delta U_{n,m}(\mathbf{z}, 0) \exp \left[-\frac{(1 - \lambda_{n,m})t}{\tau} \right], \quad (18)$$

where $\delta U_{n,m}(\mathbf{z}, 0)$ is the initial value of the projection. $u_{1,0}$ and $u_{0,1}$ correspond to the trackability of the synaptic input as well as the neuronal response, because $u_{1,0} \sim \partial \tilde{U} / \partial x_1$ and $u_{0,1} \sim \partial \tilde{U} / \partial x_2$. $u_{1,0}$ and $u_{0,1}$ are the modes of the position shift of the synaptic input. Thus, it guarantees the stability of the stationary solution and trackability of the synaptic input.

We are now ready to find the tracking solution to the stimulus abruptly changed from $(0, 0)$ to $(z_{01}, 0)$ at $t = 0$. Neglecting the dependence on all $a_{n,m}$ terms in Eq. (15), we have

$$\frac{dz_1}{dt} = \frac{\alpha}{\tau} (z_{01} - z_1) \exp \left[-\frac{(z_{01} - z_1)^2}{8a^2} \right], \quad (19)$$

which is consistent with the result obtained from the one-dimensional case [13, 14]. This approximation is useful when $|\mathbf{z} - \mathbf{z}_0|$ is small and the stimulus is weak. It also shows that the tracking behavior is similar to the one-dimensional case.

3.2. Using Basis Functions in Polar Coordinates

Since the system is also rotationally invariant, the analysis can proceed by using the eigenfunctions with polar coordinates. The eigenfunctions of quantum harmonic oscillators are

$$\psi_{l,j}(r, \theta) = \sqrt{\frac{(\frac{l-j}{2})!(\frac{l+j}{2})!}{2\pi a}} \left[\sum_{t=0}^{\frac{l-|j|}{2}} \frac{(-1)^t (\frac{r}{\sqrt{2a}})^t}{(\frac{l-j}{2})!(\frac{l+j}{2})!t!} \right] e^{-\frac{r^2}{4a^2} + i j \theta}, \quad (20)$$

where $i = \sqrt{-1}$, and l, j are the radial and angular quantum numbers respectively. Decomposing the distortion term, we have $\delta U = \sum_{l,j} b_{l,j}(t) \psi_{l,j}$. The matrix elements of the transformation matrix are

$$[T]_{n,m,l,j} \equiv \langle u_{n,m} | \psi_{l,j} \rangle = i^m \sqrt{\frac{(\frac{l+j}{2})!(\frac{l-j}{2})!}{n!m!2^l}} \left[\sum_t (-1)^t \binom{m}{t} \binom{n}{\frac{l-j}{2} - t} \right], \quad (21)$$

Similar to Eq. (10), there is an interaction kernel \mathcal{G} that represents the interaction between different $\psi_{l,j}$, which can be obtained by $\mathcal{G} = T^{-1} \mathcal{F} T$. The first few eigenfunctions of \mathcal{G} are

$$\Psi_{00} = \psi_{00}, \quad (22)$$

$$\Psi_{1\pm 1} = \psi_{1\pm 1}, \quad (23)$$

$$\Psi_{20} = \frac{1}{\sqrt{1 + (\lambda_{00} - 1/2)^2}} [\psi_{00} + (\lambda_{00} - 1/2) \psi_{20}], \quad (24)$$

$$\Psi_{2\pm 2} = \psi_{2\pm 2}, \quad (25)$$

$$\Psi_{3\pm 1} = \frac{1}{\sqrt{3^2 + 2}} [\sqrt{2} \psi_{1\pm 1} + 3 \psi_{3\pm 1}], \text{ and} \quad (26)$$

$$\Psi_{3\pm 3} = \psi_{3\pm 3}, \quad (27)$$

where the indices l and j of $\Psi_{l,j}$ represent the highest basis function it contains. Their eigenvalues are $\lambda_{00}, 1, 1/2, 1/2, 1/4$, and $1/4$ respectively.

As shown in Figs. 3 to 8, the eigenfunctions are symmetric with respect to the origin. The eigenfunctions correspond to different modes of the distortion of the synaptic input during the

motion. Ψ_{00} corresponds to the change in height. It can describe, say, the reduction of the bump height during the process to catch up with the new position of the stimulus, as shown in figure 2. $\Psi_{1,\pm 1}$ can describe the movement of the bump towards the preferred positions. $\Psi_{2,0}$ describes not only changes in the height of the bump, but also changes in the width of the bump during the motion. $\Psi_{3,\pm 1}$ describes the skewing of the bump due to the stimulus and other modes. While the above distortion modes are apparently extensions of those in the one-dimension case, the modes $\Psi_{2,\pm 2}$ and $\Psi_{3,\pm 3}$ are unique to the two-dimensional case. The former corresponds to an elliptical distortion of the bump shape, and the latter to a three-fold distortion.

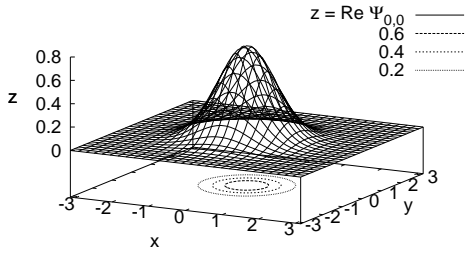


Figure 3. Real part of the eigenfunction $\Psi_{0,0}$.

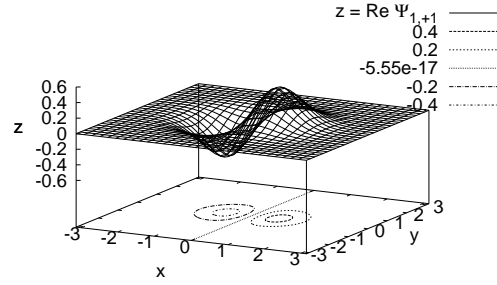


Figure 4. Real part of the eigenfunction $\Psi_{1,+1}$.

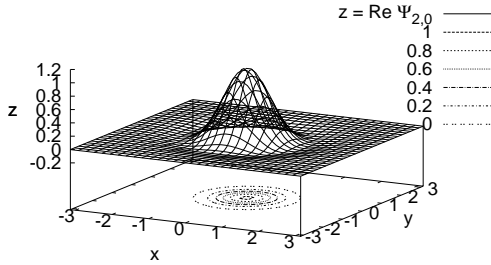


Figure 5. Real part of the eigenfunction $\Psi_{2,0}$.

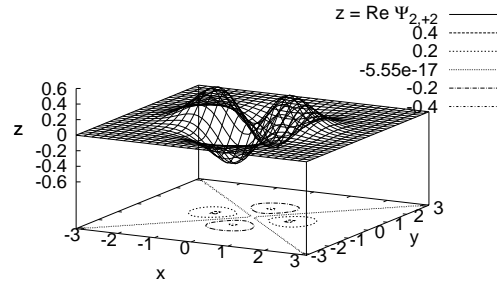


Figure 6. Real part of the eigenfunction $\Psi_{3,+1}$.

By using the transformation matrix in Eq. (21), Eq. (10) can be transformed from rectangular to polar coordinates up to arbitrary order. For the perturbation up to $l = 2$, and for external stimuli symmetric with respect to the x -axis, we have

$$\left(\tau \frac{d}{dt} + 1 - \lambda_{00}\right) b_{00} = I_{00} - \tau \frac{dR}{dt} \left[-\frac{1}{\sqrt{8}a} (b_{1-1} + b_{1+1}) \right] - b_{20}, \quad (28)$$

$$\tau \frac{d}{dt} b_{1\pm 1} = I_{1\pm 1} - \tau \frac{dR}{dt} \left(\frac{\sqrt{2}\pi a U_0 + b_{00}}{2\sqrt{2}a} - \frac{1}{2\sqrt{2}a} b_{20} - \frac{1}{2a} b_{2\pm 2} \right), \quad (29)$$

$$\left(\tau \frac{d}{dt} + \frac{1}{2}\right) b_{20} = I_{20} - \tau \frac{dR}{dt} \left(-\frac{1}{2\sqrt{2}a} (b_{1-1} + b_{1+1}) \right), \quad (30)$$

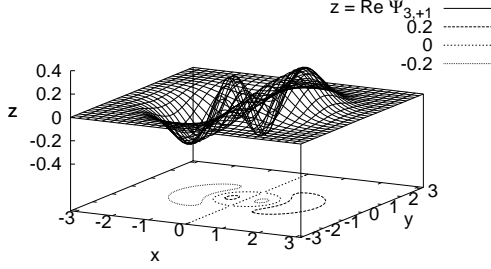


Figure 7. Real part of the eigenfunction $\Psi_{2,+2}$.

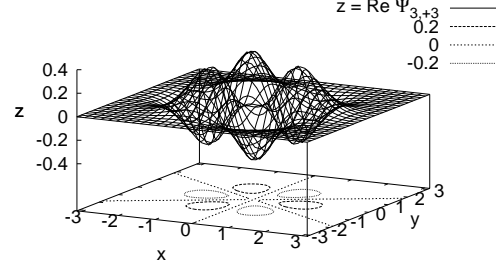


Figure 8. Real part of the eigenfunction $\Psi_{3,+3}$.

$$\left(\tau \frac{d}{dt} + \frac{1}{2}\right) b_{2\pm 2} = I_{2\pm 2} - \tau \frac{dR}{dt} \frac{1}{2a} b_{1\pm 1}, \quad (31)$$

$$\tau \frac{dR}{dt} = 2\sqrt{2}a \frac{I_{11}}{\sqrt{2\pi}aU_0 + b_{00} - \sqrt{2}b_{20} - 2\sqrt{2}b_{22}}, \quad (32)$$

where R is the radial distance from the origin. Note that $I_{l,j} = I_{l,-j}$ and $b_{l,j} = b_{l,-j}$ due to the symmetry when the stimulus lies on the x -axis.

4. Simulation Experiments

In the simulation experiments, the number of neurons is $N = N_x \times N_y$, and the range of (x_1, x_2) is $-\pi \leq x_1, x_2 < \pi$. The boundary condition is periodic.

4.1. Reaction Time to an Abrupt Change of the Stimulus

In this experiment, the stimulus is centered at $(0,0)$ until the synaptic input $U(\mathbf{x},t)$ becomes steady. At $t = 0$, the stimulus abruptly changes from $(0,0)$ to $\mathbf{z}_0 \equiv (z_{01}, 0)$, and we observe the dependence of the reaction time on the distance z_{01} . Then, the bump will track the stimulus, as shown in figure 2. The reaction time is defined by the time needed to have $|\mathbf{z}(t) - \mathbf{z}_0| < \Theta$, where Θ is the threshold. This threshold is necessary in this experiment, because the motion of $\mathbf{z}(t)$ will become very slow when it approaches the stimulus, as implied by Eq. (19). Also, the assumption is reasonable because in real biological systems, we do not need to have $\mathbf{z}(t) = \mathbf{z}_0(t)$ to make decisions.

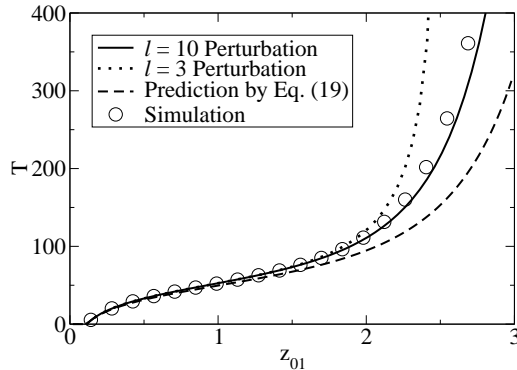


Figure 9. The reaction time T for the synaptic input to catch up the stimulus position change from 0 to z_{01} . Parameters: $N = N_x \times N_y = 40 \times 40$, $\Theta = \pi\sqrt{2/N}$, $\alpha = 0.05$ and the rest are the same as figure 1.

As shown in figure 9, the prediction given by Eq. (19) works well only when the change in the position of the stimulus is small, while the $l = 3$ perturbation works well up to $|\mathbf{z}_0| = 2$. The prediction of the $l = 10$ perturbation is the best among the three. From this result, one can state that, when the position change of the stimulus is small, only $\psi_{l,j}$ with small l will be activated. However, when the change in stimulus position is larger, higher order distortions are activated. This is reasonable, because, for smaller l , the distortions are concentrated around $\mathbf{z}(t)$, but if the stimulus is far away from $\mathbf{z}(t)$, the tail part of the bump will be distorted first, leading to higher order distortions.

4.2. Amplitudes of the Basis Function Distortion Modes

From Eqs. (10) and (21), $b_{l,j}$ can be predicted by the projections of δU onto the basis functions $\psi_{l,j}$. The experimental settings are the same as above, but \mathbf{z}_0 was fixed to be $\mathbf{z}_0 = (2, 0)$ in polar coordinates.

As shown in figure 10, the predicted $b_{l,j}$'s agree with the simulation results well. It confirms that the perturbative method can also predict the motion of the synaptic input in detail. $b_{0,0}$ indicates that the height drops from its initial value after the stimulus is shifted. The $\psi_{0,0}$ component of the distortion is reduced by the inhibition. It approaches 0 roughly, as if there were no external stimulus, and there is even a slight overshoot. Afterwards, the distortion relaxes smoothly to the equilibrium value when it approaches the shifted position of the stimulus. Similarly, the amplitude of the $\psi_{1,\pm 1}$ components falls abruptly to a negative value initially. This is due to the tail of the bump being pulled by the newly positioned stimulus, causing the peak to lag behind the center of mass. Afterwards, it relaxes smoothly to 0. The initial change in $b_{2,0}$ indicates an increase in width along the direction of the stimulus, and that of $b_{2,\pm 2}$ signals a cigar-shaped distortion when the bump is being pulled by the stimulus. The amplitude $b_{3,\pm 1}$ describes the skewness of the bump, and $b_{3,\pm 3}$ describes the bump being distorted when its tip is pulled by the stimulus, with the posterior part lagging in motion, causing a triangular-shaped distortion.

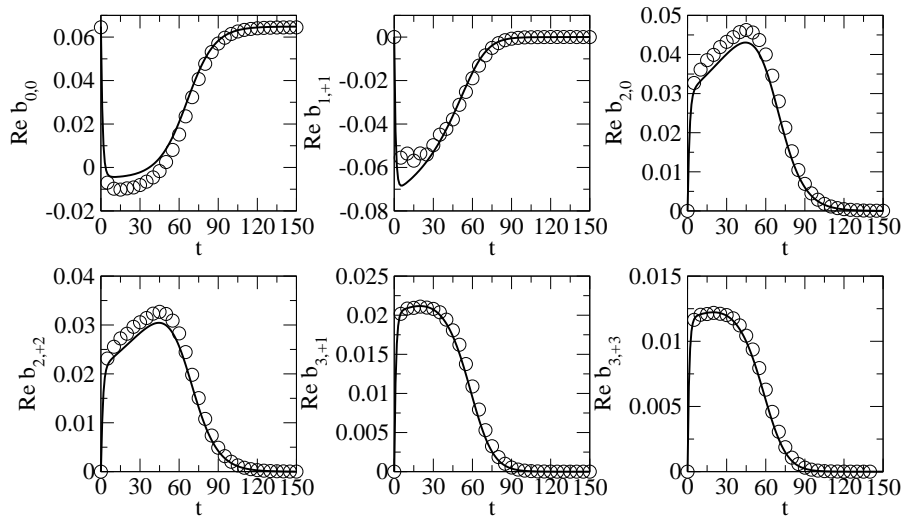


Figure 10. The experimentally projected $b_{l,j}$ and the predicted $b_{l,j}$ using $l = 10$ perturbation. Solid lines: corresponding predictions. Circles: simulation results. Parameters: same as figure 9.

5. Conclusion and Discussion

In this paper, a perturbative method to deal with continuous attractor neural networks is presented in two-dimensional space. We have introduced a simple solvable model to demonstrate how to use perturbative method to analyze the dynamics of distortions and synaptic input. Since the coupling factors are defined to be translational invariant, a family of stationary states can be sustained anywhere in the preferred stimulus space. Furthermore, the synaptic input is able to track the stimulus in the space. By studying the dynamics, one can deduce the tracking time and the distortions of the synaptic input.

For one-dimensional continuous attractor neural networks, the tracking speed can be roughly approximated by

$$\frac{dz}{dt} = \frac{\alpha}{\tau}(z_0 - z) \exp \left[-\frac{|z_0 - z|^2}{8a^2} \right], \quad (33)$$

which is similar to Eq. (19). Although they are rough approximations, the basic properties of tracking in one and two dimensions are similar. The key predictions in the one-dimensional case, the maximum trackable speed [11, 13] and the lag behind a continuously moving stimulus, are also applicable in two dimensions. These similarities arise from the rotational invariance of the interaction kernel and the bump in the limit of weak stimulus, since the description of the dynamics by the translational mode is sufficient. Note, However, for the stronger stimuli, the dynamics is richer in the two-dimensional case, since distortion modes unique to the two-dimensional case need to be considered; examples of cigar-shaped and triangular-shaped distortions are shown in figure 10.

For this particular model, the eigenfunctions of the interaction matrix in polar coordinates are also studied. By Eq. (21), the interaction matrix \mathcal{G} can be obtained. However, due to complications in the calculation, we can only calculate it term by term. Since the matrix is upper triangular, the eigenvalues are the diagonal entries, $\lambda_{n,m}$. As the eigenvalues are at most 1, one can show that the synaptic input has a stable Gaussian form. The eigenfunctions corresponding to eigenvalue 1 represent the positional shift. As studied above, different modes of distortion correspond to different kinds of distortion. For example, the component of $\psi_{0,0}$ corresponds to the change in height, while $\psi_{2,0}$ represents the change in width.

Concerning the robustness of the method, we remark that it can be applied to other types of networks with tracking behavior. A common example is the continuous attractor neural network with the Mexican hat interaction. Using the basis functions of the two-dimensional quantum harmonic oscillator, we can obtain the matrix elements of the interaction kernel numerically, although elegant expressions such as those obtained here may not be available. Perturbation dynamics can then be worked out analogously. This proposed extension can be able to address a recent issue of interest, namely, the instability of bumps and rings in a two-dimensional neural field of Amari type [19]. Meanwhile, we note in passing that the present model with a quadratic response and a global inhibition does not suffer from the stability problem.

Acknowledgement

This work is partially supported by the Research Grants Council of Hong Kong (Grant Nos. HKUST 603607 and HKUST 604008).

References

- [1] Amari S 1977 *Biol. Cybern.* **27** 77–87
- [2] Ermentrout B 1998 *Rep. Prog. Phys.* **61** 353–430
- [3] Seung H S 1996 *Proc. Acad. Sci. USA* **93** 13339–44
- [4] Wu S, Hamaguchi K and Amari S 2008 *Neural Comput.* **20** 994–1025
- [5] Zhang K C 1996 *J. Neurosci.* **16** 2112–26
- [6] Ben-Yishai R, Lev Bar-Or R and Sompolinsky H 1995 *Proc. Natl. Acad. Sci. USA* **92** 3844–8

- [7] Georgopoulos A P, Taira M and Lukashin A 1993 *Science* **260** 47–52
- [8] Samsonovich A and McNaughton B L 1997 *J. Neurosci.* **7** 5900–20
- [9] Jastorff J, Kourtzi Z and Giese M 2006 *J. Vision* **6** 791
- [10] Folias E and Bressloff P 2004 *SIAM J. Appl. Dyn. Syst.* **3** 378–407
- [11] Hansel D and Sompolinsky H 1998 *Methods in Neuronal Modeling: From Ions to Networks* ed Koch C and Segev I (MIT Press, Cambridge)
- [12] Wu S and Amari S 2005 *Neural Comput.* **17** 2215–39
- [13] Fung C C A, Wong K Y M and Wu S 2008 *Europhys. Lett.* **84** 18002
- [14] Fung C C A, Wong K Y M and Wu S *Neural Comput.* (*in press*)
- [15] Coombes S and Owen M R 2005 *Phys. Rev. Lett.* **94** 148102
- [16] Taylor J G 1999 *Biol. Cybern.* **80** 393–409
- [17] Werner H and Richter T 2001 *Biol. Cybern.* **85** 211–217
- [18] Griffiths D 2004 *Introduction to Quantum Mechanics* (Prentice Hall)
- [19] Owen M R, Liang C R and Coombes S 2007 *New J. Phys.* **9** 378